AFRL-HE-AZ-TP-2002-0002

UNITED STATES AIR FORCE RESEARCH LABORATORY

KNOWLEDGE ACQUISITION IN DISTRIBUTED MISSION TRAINING

Roger Schvaneveldt
Arizona State University
Tempe AZ 85287

Richard Tucker
SPC/TACCSF
Kirtland AFB NM 87117

Anna Rae Castillo
Winston Bennett Jr

Air Force Research Laboratory
Warfighter Training Research Division
6030 South Kent Street
Mesa AZ 85212-6061

December 2002

AIR FORCE MATERIEL COMMAND
AIR FORCE RESEARCH LABORATORY
HUMAN EFFECTIVENESS DIRECTORATE
Warfighter Training Research Division
6030 South Kent Street
Mesa AZ 85212-6061

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)
01-12-2002

2. REPORT TYPE
Interim

3. DATES COVERED (FROM - TO)
xx-11-2000 to xx-11-2001

4. TITLE AND SUBTITLE
Knowledge Acquisition in Distributed Mission Training
Unclassified

5a. CONTRACT NUMBER
F41624-97-D-5000

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62205F

6. AUTHOR(S)
Schbaneveldt, Roger ; Author
Tucker, Richard ; Author
Castillo, Anna ; Author
Bennett, Winston Jr. ; Author

5d. PROJECT NUMBER
1123

5e. TASK NUMBER
B0

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME AND ADDRESS
Arizona State University
Tempe, AZ85287

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME AND ADDRESS
Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division
Mesa, AZ85212-6061

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL; AFRL/HEA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-HE-AZ-TP-2002-0002

12. DISTRIBUTION/AVAILABILITY STATEMENT
APUBLIC RELEASE
,

13. SUPPLEMENTARY NOTES
Air Force Research Laboratory Technical Monitor: Dr Winston Bennett Jr., 480-988-6561 X-297, DSN 474-6297 This paper was presented at the Interservice/Industry Training Systems and Education Conference, held in Orlando FL, Dec 2001.

14. ABSTRACT
The Air Force Research Laboratory (AFRL) in Mesa, AZ, has an ongoing program of research on training in a 4-ship F-16 Distributed Mission Training (DMT) system. A team of four pilots comes to the laboratory for a week-long training exercise and fly together as a 4-ship team on several missions designed to provide exposure to a range of combat scenarios. Missions involve briefing and debriefing sessions in addition to simulator time. Several projects are underway to assess the training effectiveness and compare different training methods. Pilots rated the relatedness of all pairs of 21 concepts from the domain of air-to-air engagements both before and after training. Measures of the internal consistency (Coherence) of the ratings and Pathfinder networks were derived from the ratings. Data from a group of the most experienced pilots (experts) provided a point of reference for the less-experienced pilots. There was significant correlation between experience level and similarity to the expert reference group at the beginning of the week supporting the general validity of the measurement methods. There was a significant negative correlation between experience level and change in similarity to experts from pre- to post training ratings. There was a significant difference in mean Coherence between the experts and novices at the beginning of the week but not at the end of the week. These measurement methods appear to provide a basis for evaluating conceptual change. These assessment methods should prove to be a useful adjunct to performance-based methods of assessing training.

15. SUBJECT TERMS
Distributed mission training; DMT; Knowledge acquisition; Performance-based methods; Performance measurement; Training assessment; Training effectiveness; Training methods; Training research;

16. SECURITY CLASSIFICATION OF:

a. REPORT	b. ABSTRACT	c. THIS PAGE
Unclassified	Unclassified	Unclassified

17. LIMITATION OF ABSTRACT
Same as Report (SAR)

18. NUMBER OF PAGES
19

19. NAME OF RESPONSIBLE PERSON
Casey, Liz
liz.casey@williams.af.mil

19b. TELEPHONE NUMBER
International Area Code
Area Code Telephone Number
480988-6561
DSN
474-6188

NOTICES

Publication of this paper does not constitute approval or disapproval of the ideas or findings. It is published in the interest of STINFO exchange.

Using Government drawings, specifications, or other data included in this document for any purpose other than Government-related procurement does not in any way obligate the US Government. The fact that the Government formulated or supplied the drawings, specifications, or other data, does not license the holder or any other person or corporation, or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

WINSTON BENNETT JR
Project Scientist

HERBERT H. BELL
Technical Advisor

CURTIS J. PAPKE, Colonel, USAF
Chief, Warfighter Training Research Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
December 2002	Interim	November 2000 to November 2001

4. TITLE AND SUBTITLE

Knowledge Acquisition in Distributed Mission Training

5a. CONTRACT NUMBER F41624-97-D-5000

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER 62205F

6. AUTHOR(S)

Roger Schvaneveldt
Richard Tucker
Anna Castillo
Winston Bennett Jr.

5d. PROJECT NUMBER 1123

5e. TASK NUMBER B0

5f. WORK UNIT NUMBER 01

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Arizona State University
Tempe AZ 85287

SPC/TACCSF
Kirtland AFB NM 87117

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division
6030 South Kent Street
Mesa AZ 85212-6061

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division
6030 South Kent Street
Mesa AZ 85212-6061

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL; AFRL/HEA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-HE-AZ-TP-2002-0002

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Air Force Research Laboratory Technical Monitor: Dr Winston Bennett Jr., 480-988-6561 X-297, DSN 474-6297
This paper was presented at the Interservice/Industry Training Systems and Education Conference, held in Orlando FL, Dec 2001.

14. ABSTRACT

The Air Force Research Laboratory (AFRL) in Mesa, AZ, has an ongoing program of research on training in a 4-ship F-16 Distributed Mission Training (DMT) system. A team of four pilots comes to the laboratory for a week-long training exercise and fly together as a 4-ship team on several missions designed to provide exposure to a range of combat scenarios. Missions involve briefing and debriefing sessions in addition to simulator time. Several projects are underway to assess the training effectiveness and compare different training methods. Pilots rated the relatedness of all pairs of 21 concepts from the domain of air-to-air engagements both before and after training. Measures of the internal consistency (Coherence) of the ratings and Pathfinder networks were derived from the ratings. Data from a group of the most experienced pilots (experts) provided a point of reference for the less-experienced pilots. There was significant correlation between experience level and similarity to the expert reference group at the beginning of the week supporting the general validity of the measurement methods. There was a significant negative correlation between experience level and change in similarity to experts from pre- to post training ratings. There was a significant difference in mean Coherence between the experts and novices at the beginning of the week but not at the end of the week. These measurement methods appear to provide a basis for evaluating conceptual change. These assessment methods should prove to be a useful adjunct to performance-based methods of assessing training.

15. SUBJECT TERMS

Distributed mission training; DMT; Knowledge acquisition; Performance-based methods; Performance measurement; Training assessment; Training effectiveness; Training methods; Training research;

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Ms Liz Casey
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED	19	480.988.6561 x- 188 DSN 474-6188

KNOWLEDGE ACQUISITION IN DISTRIBUTED MISSION TRAINING

Roger Schvaneveldt
Arizona State University
Mesa, AZ

Richard Tucker
SRC/TACCSF
Albuquerque, NM

Anna Castillo
Air Force Research Lab
Mesa, AZ

Winston Bennett, Jr.
Air Force Research Lab
Mesa, AZ

SUMMARY

At the Air Force Research Laboratory (AFRL) in Mesa, AZ, there is an ongoing program of research on training in a 4-ship F-16 Distributed Mission Training (DMT) system. Typically, a team of four pilots comes to the laboratory for a week-long training exercise. They fly together as a 4-ship team on several missions designed to provide exposure to a range of combat scenarios. The missions involve extensive briefing and debriefing sessions in addition to the time in the simulators. In order to track the effects of training, several inter-related projects are underway to assess the effectiveness of training and to compare different training methods. In the project reported here, we assessed changes in the ways pilots understand important concepts related to the training. Pilots rated the relatedness of all pairs of 21 concepts from the domain of air-to-air engagements both before and after training. Measures of the internal consistency (Coherence) of the ratings and Pathfinder networks were derived from the ratings. Data from a group of the most experienced pilots (experts) provided a point of reference for the less-experienced pilots. At the beginning of the week, Coherence was significantly correlated with previous experience in fighter aircraft suggesting that providing consistent ratings depends on having a well-developed mental model of the domain. Also, there was a significant correlation between experience level and similarity to the expert reference group at the beginning of the week supporting the general validity of the measurement methods. There was a significant negative correlation between experience level and change in similarity to experts from pre- to post training ratings. Greater changes were found for the least-experienced pilots. As a result of these changes, correlations with prior experience level were no longer statistically significant at the end of the week. Further analyses on a group of the least-experience pilots (novices) lead to similar conclusions. In particular, there was a significant difference in mean Coherence between the experts and novices at the beginning of the week but not at the end of the week. Also novices showed significant pre- to post training increases in Coherence and in similarity to experts. These measurement methods appear to provide a basis for evaluating conceptual change. These assessment methods should prove to be a useful adjunct to performance-based methods of assessing training.

Roger Schvaneveldt is a Professor of Applied Psychology at Arizona State University, East in Mesa, AZ. He earned the Ph.D. in psychology at the University of Wisconsin in 1967. His research interests include basic issues in memory and cognition as well as applied research in human factors, in general, and aviation, in particular. He developed the Pathfinder network scaling used in the present investigation.

Richard Tucker is Program Manager for Scientific Research Corporation at the Theater Air Command and Control Simulation Facility, (TACCSF) Kirtland AFB, NM. He is a retired Air Force Colonel and Fighter Pilot who received an MS in Psychology at the University of South Dakota in 1977. His interests focus on realistic training through the use of Distributed Mission Training and Distributed Mission Operations.

Anna Castillo is the Link Simulation and Training team lead for Distributed Mission Training research at the Air Force Research Laboratory. Her research objectives are to increase Distributed Mission Training effectiveness and to determine the extent to which lessons learned in DMT transfer to the operational unit. She received the MS in experimental psychology from New Mexico Highlands University in 1998.

Winston Bennett, Jr. is team leader for combat mission analysis, rehearsal, training technology and performance assessment R&D at AFRL. He holds a Ph.D. in Industrial Psychology from Texas A&M University. Honors include the 1998 American Society for Training and Development (ASTD) Research Award and the 1999 AFRL Harry G. Armstrong Scientific Achievement Award.

KNOWLEDGE ACQUISITION IN DISTRIBUTED MISSION TRAINING

Roger Schvaneveldt
Arizona State University
Mesa, AZ

Richard Tucker
SRC/TACCSF
Albuquerque, NM

Anna Castillo
Air Force Research Lab
Mesa, AZ

Winston Bennett, Jr.
Air Force Research Lab
Mesa, AZ

INTRODUCTION

F-16 DMT Training

At the Air Force Research Laboratory (AFRL) in Mesa, AZ, there is an ongoing program of research on training in a 4-ship F-16 Distributed Mission Training (DMT) system. Typically, a team of four pilots comes to the laboratory for a week-long training exercise. They fly together as a 4-ship team on several missions designed to provide exposure to a range of combat scenarios. The missions involve extensive briefing and debriefing sessions in addition to the time in the simulators. The simulators at AFRL can be connected in real-time to simulators at other locations in the United States and to stations involving other roles in addition to fighter pilots such as AWACS operators. Thus, a variety of training settings can be devised in the overall DMT system. Several research projects are focused on designing and evaluating training. In this paper, we discuss one line of research on evaluating fighter-pilot training.

Assessment of Training

There are many reasons to be concerned with evaluating the training provided by this system. Of primary importance is to determine that the objectives of the training are being met. There is also concern with developing training methods that will produce the largest gains possible from the week of training. Numerous research questions can be framed concerning performance in the simulators, performance in actual flight, and the acquisition of requisite knowledge.

Simulator performance. Perhaps the most direct approach to evaluating the impact of training is to assess performance in the simulated missions encountered as part of training. Such efforts are underway in the lab, but there are some problems associated with using performance in the simulators as the basis for evaluating progress in training. First, because the complexity of the missions typically increases dramatically over the course of the week, direct comparisons of performance are suspect. Using benchmark missions has been explored, but finding an appropriate level of difficulty for such missions is problematic because using more complex missions early in the week may place unrealistic demands on pilots who need additional experience. Conversely, easier missions may not be sufficiently challenging to show deficiencies even earlier in the week.

Performance in missions or further training. Ultimately, training should have a positive impact on performance in actual missions flown and on the progress of further training in operational units. Efforts are also underway to gather and analyze data that bear on the transfer of training to actual flying performance and on performance in further training. Such data take considerable time and effort to collect, and there are major issues associated with developing appropriate methods and measurements to assess performance in flight. In this project, we concentrated on evaluating pilots' acquisition of the knowledge required to understand combat situations and to make good decisions about possible courses of action.

Knowledge. Still another approach to evaluating the impact of training is to examine changes in knowledge as a function of training. Clearly, the ultimate goal of pilot training is to produce expertise in performance in flight. However, it is also clear that expert performance depends on the acquisition of both knowledge and skill. Assessing skill requires observation in live performance, but assessing knowledge can be done outside the live performance environment. This independence of assessment from the training allows greater control over the collection of data pertinent to assessing knowledge. In this investigation, we choose to approach knowledge by means of how pilots see the relationships among various concepts in the air-to-air combat maneuvering domain. This area receives considerable attention in the simulator training so it is natural to suppose that there should be an impact on pilot's understanding of these concepts.

The technique we used is generally known as Pathfinder Network Scaling (see Schvaneveldt, Dearholt, & Durso, 1989; Schvaneveldt, Durso, & Dearholt, 1989; Schvaneveldt, 1990). The method uses individuals' judgments as a source from which to extract underlying network structures. Among various applications, this method has been used to capture expert-novice differences in conceptual structure (USAF pilots – Schvaneveldt, et al., 1985; also see Cooke & Schvaneveldt, 1988 and Rowe, Cooke, Hall & Halgren, 1996), to assess student knowledge acquisition (Goldsmith, Johnson, & Acton, 1991 and to analyze and design user-system interfaces (McDonald & Schvaneveldt, 1988). Pathfinder networks stem from research and theory in cognitive psychology, and they represent one example of a bridge between basic research and application (see Gillan & Schvaneveldt, 1999).

Graph Theory and Networks

The abstract model underlying the network model of relationships among concepts is graph and network structures. In mathematical graph theory, a graph is an abstraction consisting of a set of nodes and a set of pairs of the nodes (Harary, 1969). Each such pair of nodes is called a link. The links can be directed or undirected. The nodes represent some entities (concepts in this case.) and the links represent relationships between nodes.

A network is a graph with weights (or costs) associated with the links. An example of a network using familiar concepts from aviation is shown in Figure 1 (from Schvaneveldt, et. al, 2000). We often expand these basic definitions to include distinguishing different types of nodes and different types of links. Taken as a whole the collection of nodes and links can represent how individuals or groups of individuals view the relationships among concepts. As in previous research, we hypothesize that differences in expertise will be reflected by differences in link structures.

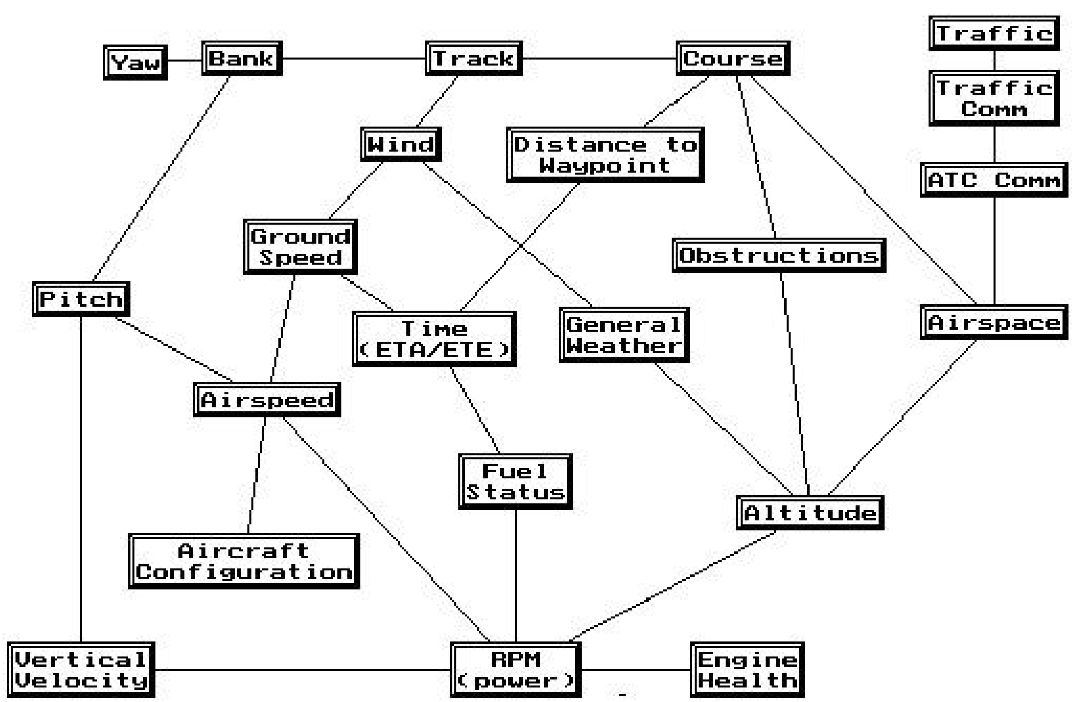

Figure 1. A Pathfinder Network for Selected Aviation Concepts.

METHOD

Participants

In all, rating data were obtained from 84 Air Force pilots. These pilots came from several different operational units around the United States. They varied greatly in experience, and they came to the laboratory with different objectives in mind. Several analyses of the data were based on all of these pilots. Some focused analyses were performed on two groups of pilots, a group of six pilots with the highest levels of experience (more than 1900 hours of flying fighter aircraft) and another group of eight with the least experience (80-105 hours of flying fighter aircraft).

Concept Selection

The concepts rated by the pilots were selected from advanced air-to-air combat maneuvering scenarios (see Table 1).

Table 1. Concepts Rated for Relatedness

Alamo	AMRAMM
Bandit/Hostile	Beam deploy
BVR	Doppler pick
Factor bandit range	Grinder
IRMD	Launch and leave
MOR	Multiple groups in azimuth
Notch	Multiple groups in range
PID	Pit Bull
Preserve range	Real world ROE
Targeting/Sorting	Unknown group
Visual mutual support	

Concept Rating

Pathfinder scaling requires measures of proximity for all pairs of elements to be scaled. In this study, ratings were based on 210 judgments of the degree of relatedness for the concepts in each pair of the 21 concepts. From such data, the Pathfinder scaling algorithm produces a network showing the connections between the various concepts.

Variables

In this section, we define the variables investigated in the analyses of the data.

We used the logarithm of the number of hours flying F-16 fighter aircraft as the measure of *Experience*.

Aside from measures of pilot experience, all other variables in the data analysis were based on the pilots' ratings of all the pairs of concepts. All 84 pilots performed the rating task at the beginning of the week (*Rating1*), but only 60 completed the ratings at the end of the week (*Rating2*). A network, PFNET(r = infinity, q = 2), was derived from each set of ratings. The networks are referred to as *Net1* and *Net2* for *Rating1* and *Rating2*, respectively.

Expert reference data were obtained by averaging the ratings given by the six most experienced pilots. We refer to this set of average ratings as *AveExpRate*. Applying the Pathfinder scaling algorithm (r = infinity, q = 2) to *AveExpRate* produced the referent expert PFNET, the *ExpNet*.
One additional analysis of rating data yields a coherence measure which assesses the internal consistency of the ratings. Coherence is computed in two steps. First, a correlation of ratings for each pair of items is computed. For example for items 1 and 2, the ratings of item I across all of the other items is correlated with item 2 across all of the other items. This determines the extent to which the items in a pair are rated similarly against the other items. The second step correlates these correlations with the ratings given for the pairs. This can be seen as comparing the direct rating of a pair with the indirect relatedness inferred from the similarity of the ratings for the items in each pair.

The coherence measure has been shown to reflect levels of expertise in that raters with more expertise in the concept domain generally produce higher coherence scores compared with less experienced raters. Coherence can also reveal unsystematic ratings that might be provided by a participant who does not take the rating task seriously resulting in careless ratings. Given that the rating task requires a large number of judgments, it is difficult for participants to be consistent by remembering earlier ratings. Rather, consistency more likely stems from using a clear understanding of the domain as a basis of the ratings. Coherence was computed from both *Rating1* and *Rating2* yielding *Coh1* and *Coh2*, respectively. *CohChange* is *Coh2 – Coh1*.

We also defined variables to assess the extent to which the ratings and the networks obtained from individuals are similar to the measures obtained from the experts.

Rate1Sim assesses the similarity of individual's ratings at the beginning of the week with the average of the experts' ratings. It is obtained by correlating *AveExpRate* with *Rating1* for each pilot.

Rate2Sim assesses the similarity of individual's ratings at the end with the

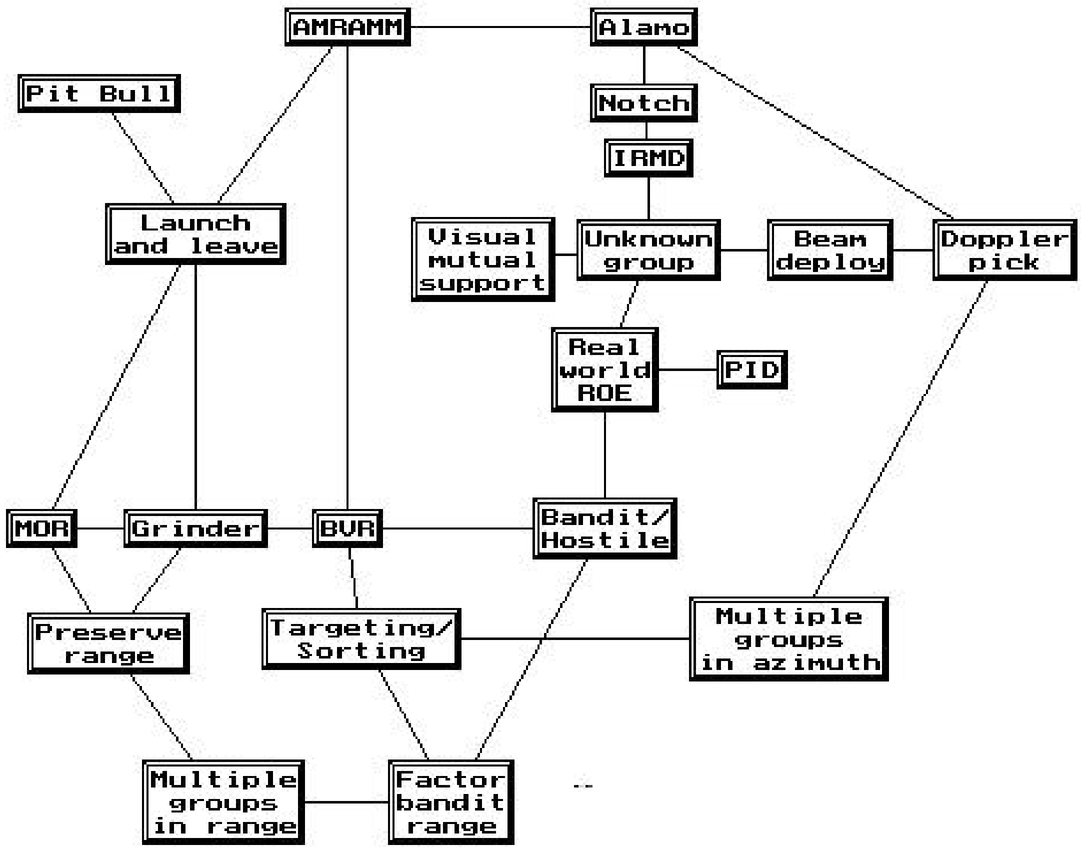

Figure 2. The Pathfinder Network Derived from the Average of Expert Ratings.

There were some indications in these data that the effects we were observing were largely due to the least experienced pilots in the sample. When these 8 pilots were removed from the data set, all of the correlations reported above became non-significant. We decided that a closer look at the data from these 8 Novice Pilots was in order (see Table 3).

Novices

It is interesting to note that all of the changes were significant for the novices as is shown by the *t*-tests at the bottom of the table. Novices show significant increases in Coherence and in their similarity to experts both in rating similarity and in network similarity.

Not shown in Table 3 is a comparison of the rating coherence between these novice pilots and the group of six experts. For the novice ratings at the beginning of the training (*Coh1*), there was a significant difference between the novices and the experts (Novice Mean = 0.34, expert Mean = 0.58, t = 2.70, df = 12. p = 0.019). After training, the difference is not significant. For these differences to appear with such a small sample size is intriguing, but the findings should be replicated with additional observations.

Overall, it appears that the methods used in the present investigation are most sensitive to the impact of training on the least experienced pilots. Perhaps it is not surprising that these pilots should show the greatest impact of the training. They certainly have a greater distance to go in achieving expert-level knowledge and performance.

Table 3. Summary of Data from Novice Pilots (<100 hours)

Coherence			Similarity of Experts and Novices Ratings			Similarity of Experts and Novices Networks		
Coh1	**Coh2**	**Coh Change**	**Rate1 Sim**	**Rate2 Sim**	**Rate Sim Change**	**Net1 Sim**	**Net2 Sim**	**Net Sim Change**
0.19	0.44	0.26	0.32	0.43	0.11	0.03	0.04	0.01
0.44	0.38	-0.06	0.56	0.63	0.07	0.11	0.15	0.04
0.36	0.54	0.18	0.37	0.63	0.25	0.07	0.10	0.04
0.50	0.69	0.20	0.50	0.62	0.12	0.05	0.05	0.01
0.59	0.57	-0.02	0.51	0.64	0.14	0.07	0.11	0.03
0.22	0.39	0.17	0.32	0.52	0.20	0.05	0.15	0.10
0.47			0.54			0.09		
-0.03	0.33	0.36	-0.11	0.48	0.58	-0.01	0.13	0.14
0.34	0.48	0.15	0.38	0.56	0.21	0.06	0.11	0.05
	t	2.78		***t***	3.17		***t***	2.78
	df	6		***df***	6		***df***	6
	p	0.032		***p***	0.019		***p***	0.032

Teams and Similarity Changes

An analysis of the similarity of pilots to one another was devoted to determining whether the members of the teams of four pilots tend to become more similar to one another over the course of the week's training in the simulators. While the data show that the ratings of team members are more highly correlated at the end of the week (*Rating2*, mean $r = .455$) than they are at the beginning of the week (*Rating1*, mean $r = .393$). The difference (.062) is significant ($p < .001$). Although it is tempting to see these data as supporting development of common mental models among team members, it is instructive to examine the similarity of pilots in different teams. A very similar result occurs comparing members of different teams. Pilots show higher correlations at the end of the week (mean $r = .426$) than at the beginning of the week (mean $r = .384$). This .042 difference is also significant ($p < .001$). Apparently the increase in similarity should be seen as an effect of the common training environment and the similarity of the situations encountered in training rather than changes that are peculiar to the 4-ship teams.

Identifying Discriminating Concepts

A final analysis was focused on identifying the concepts that are viewed most distinctly by expert and novice pilots. Ratings of novice pilots were averaged (*AveNovRate*). Then the differences between *AveNovRate* and *AveExpRate* were computed. Summing the absolute values of these differences for each of the 21 concepts allows us to order them by the magnitude of the differences between expert and novice ratings. This ordering is shown in Table 4.

Table 4. Concepts that Discriminate

Most Discriminating		Least Discriminating	
Rank	**Concept**	**Rank**	**Concept**
1	Grinder	1	Unknown group
2	Pit Bull	2	Alamo
3	MOR	3	Doppler pick
4	Multiple groups in range	4	Notch
5	Preserve range	5	AMRAMM
6	BVR	6	Multiple groups in azimuth
7	Beam deploy	7	PID
8	Visual mutual support	8	Real world ROE
9	Launch and leave	9	Targeting/Sorting
10	IRMD		
11	Factor bandit range		
12	Bandit/Hostile		

Knowing about the relative difficulty of these concepts could usefully feed back into the design of training by showing what the novices seem to know the least about. By ensuring that these concepts receive adequate coverage in the briefing and debriefing sessions, novices may be better prepared to deal with the corresponding aspects of the simulator flights.

SUMMARY AND CONCLUSIONS

Several aspects of the data reported here showed that a week of training in the DMT system at AFRL produces measurable changes in the way less experienced pilots view a collection of concepts from the domain of air-to-air combat. The major effects were found for the least experienced pilots in our sample. Perhaps pilots with more experience have already mastered the understanding of the domain so training and practice does not lead to major changes. It is possible that another set of more advanced concepts would reveal changes in more experienced pilots, and there may be other methods of assessment that should be pursued.

We are developing a method for more directly evaluating knowledge associated with assessing combat situations and selecting appropriate courses of action. Watch for further developments.

Although there were overall increases in similarity among pilots after training, there was no indication that member of teams become more similar than do members of different teams. The increased similarity is apparently due to pilots being exposed to similar scenarios in training rather than to pilots learning more about members of pilots' teams. Other studies of team cognition (see Cooke, Salas, Cannon-Bowers, & Stout, 2000) have ofter looked at increasing similarity amont team members. The data reported here suggest that it may be useful to determine the extent to which increasing similarity is due to exposure to similar task conditions as opposed to training experiences with members of a team.

The results of our study have had some effects on the design of the training environment. The concepts we identified as leading to the greatest differences between novices and experts have received more attention in the briefing and debriefing sessions accompanying the training sessions. Other ways in which the information we collect can be used in training are being considered. For example, pilots could be presented with networks like the one shown in Figure 2 to elicit discussion about the nature of the relationships depicted by the links. Inevitably such discussion draws out similarities and differences in the ways pilots think about the concepts. A hypothesis that could be investigated is that training would benefit from such discussions, discussions that are driven by pilots underlying conception of air-to-air combat.

We are also pursuing ways of tapping into more procedural knowledge using rating and network scaling procedures. There is much more to be done.

REFERENCES

Cooke, N. M., & Schvaneveldt, R. W. (1988). Effects of computer programming experience on network representations of abstract programming concepts. *International Journal of Man-Machine Studies, 29*, 407-427.

Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. (2000). *Measuring team knowledge*. Human Factors, 42, 151-173.

Gillan, D. J., & Schvaneveldt, R. W. (1999). Applying cognitive psychology: Bridging the gulf between basic research and cognitive artifacts. In F. Durso, R. Nickerson, R. Schvaneveldt, S. Dumais, M. Chi, & S. Lindsay (Eds.), *The handbook of applied cognition*. (pp. 3-31) Chichester: Wiley.

Goldsmith, T. E., Johnson, P. J., Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology*, 83, 88-96.

Harary, F. (1969). *Graph theory*. Reading, MA: Addison-Wesley.

McDonald, J. E., & Schvaneveldt, R. W. (1988). The application of user knowledge to interface design. In R. Guindon (Ed.), *Cognitive Science and Its Applications for Human-Computer Interaction*, Hillsdale: Erlbaum.

Rowe, A. L., Cooke, N. J., Hall, E. P. & Halgren, T. L. (1996). Toward an online assessment methodology: Building on the relationship between knowing and doing. *Journal of Experimental Psychology: Applied*, 2, 31-47.

Schvaneveldt, R. W. (Ed.) (1990). *Pathfinder associative networks: Studies in knowledge organization.* Norwood, NJ: Ablex.

Schvaneveldt, R.W., Beringer, D., Lamonica, J., Tucker, R., & Nance, C. (2000). *Priorities, organization, and sources of information accessed* *by pilots in various phases of flight.* Final Report DOT/FAA/AM-00/26, Federal Aviation Administration

Schvaneveldt, R. W., Dearholt, D. W., & Durso, F. T. (1988). Graph theoretic foundations of Pathfinder networks. *Computers and Mathematics with Applications*, *15*, 337-345.

Schvaneveldt, R. W., Durso, F. T., & Dearholt, D. W. (1989). Network structures in proximity data. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*, Vol. 24 (pp. 249-284). New York: Academic Press.

Schvaneveldt, R. W., Durso, F. T., Goldsmith, T. E., Breen, T. J., Cooke, N. M., Tucker, R. G., & DeMaio, J. C. (1985). Measuring the structure of expertise. *International Journal of Man-Machine Studies*, *23*, 699-728.